Jurnal Ilmu Komputer dan Informasi (Journal of Computer Science and Information)
18/1 (2025), 103-111. DOI: http://dx.doi.org/10.21609/jiki.v18i1.1428

Utilizing X Sentiment Analysis to Improve Stock Price Prediction Using Bidirectional Long Short-Term Memory

Marcella Komunita Pasaribu, Fahrel Gibran Alghany, Benhard Simanullang, Hilma Nur Khasanah, Hanan Nurul Hardyana Zain, Khadijah*, and Rismiyati

Department of Informatics, Faculty of Science and Mathematics, Universitas Diponegoro, Semarang, 50275, Indonesia

E-mail: khadijah@lecturer.undip.ac.id

Abstract

The capital market is one of the important factors that influence the national economy. However, the stock price in capital market fluctuates over time. Therefore, the investors strongly need an accurate prediction of stock price for making profitable decision. However, with the pervasive influence of the internet, investors and investment institutions have started incorporating online opinions and news, including those found on social media platforms like X. This research aims to enhance stock price prediction by utilizing X sentiment analysis. The sentiment of tweets from X related to IHSG stock price is predicted by using BERT (Bidirectional Encoder Representations from Transformers), then its result is integrated with the historical stock price data for predicting future stock price by using BiLSTM (Bidirectional Long Short-Term Memory). The experiment results show that the RMSE and MAPE of the proposed model with sentiment analysis is decreased by 0.042 and 0.595, resepectively, compared to the model without sentiment analysis. Therefore, it can be concluded that the inclusion of X sentiment analysis in conjunction with BiLSTM succeeded in improving the performance of stock price prediction. The study's outcome is expected to be valuable for investors to make profitable decisions, leveraging the information available on social media.

Keywords: *BERT, BiLSTM, sentiment analysis, stock price prediction, X*

1. Introduction

The capital market is one of the important factors that influence the national economy and finance. The presence of the capital market as an alternative funding source for both government and private companies plays a crucial role. The capital market functions as a platform for collaboration between investors and issuers, while also providing opportunities for investors to earn profits [1].

In the current era of globalization and rapid technological advancement, the increased investment in the capital market in Indonesia is one of the impacts of technological development. The internet enables capital market transactions to be conducted without limitations of time and location. In capital market investment, speed and accuracy in transactions are crucial for investors to make profitable decisions.

However, various macroeconomic factors (such as inflation, interest and exchange rates) and current events have influences to the fluctuation of stock price in the capital market. Therefore, numerous studies have been conducted to predict the increase and decrease in stock prices to assist investors in decision-making. Traditional investment tools primarily relied on quantitative variables related to macroeconomic factors [2].

In addition to using economic indicators, a popular branch of Artificial Intelligence (AI), namely, Machine Learning (ML), can be utilized to perform stock price prediction. ML allows the machine to learn the pattern from the historical data to predict what happens in the future. In the case of stock price prediction, ML can be utilized to predict the future stock price so that investors can decide which stocks they should buy or sell. Previous researches had taking advantage of ML algorithms to perform stock price prediction based on historical stock price data [3-4].

Recurrent Neural Networks (RNNs) are widely used for predicting time series data, which align with the nature of historical stock price data that can be measured over time and specific intervals. However, RNNs have a limitation known as the vanishing gradient problem during recurrent iterations. To address this limitation, Long-Short Term Memory (LSTM) neural networks have been introduced to handle vanishing gradient problem in the original RNN model [5].

Previous researches have utilized LSTM for predicting time series data. Ozdemir et al. employed LSTM to create nickel price prediction model which achieved the MAPE of 7.06% [6]. Yao and Yan also perform stock price prediction for Shanghai Stock Exchange by integrating LSTM and DLWR (dynamic local weighted regression) and reached the MAPE around 1% [7]. Md et al. utilized multi-layer sequential LSTM to predict stock Price of Samsung. Their method was able to achieve MAPE of 2.18% on the testing set [8]. Those researches used historical or past dataset and LSTM algorithm to predict the future price and were able to achieved MAPE less than 10%.

Currently, there is a variation of LSTM network, namely BiLSTM (Bidirectional LSTM), which perform sequence processing in forward and backward direction. By using BiLSTM, the performance of the model is able to increase due to the value of the next hidden state will be formulated based on the previous and subsequent information [9]. BiLSTM is able to outperform unidirectional LSTM to predict edge weight in dynamic cross market equity network in selected Asian stock market indices [10].

In addition to use historical data, stock price prediction can be improved by utilizing sentiment analysis of public opinions because investors' opinions may influence market conditions. Many investors make decisions based on the latest news and express their opinions [11]. In their research, Jing et al. concluded that the stock price prediction is better when using the hybrid feature from sentiment analysis and past stock price data [12]. Liu et al. also stated form their research that there is positive synergy between sentiment of investor and stock price in market [13].

Sentiment analysis is a kind of task in Natural Language Processing (NLP) that aims to analyze and determine the polarity of a text. Sentiment analysis produces text labels indicating certain polarity, such as a negative or positive sentiment. Sentiment analysis can be conducted by using one of the three available approaches, namely machine learning, rule-based, and lexical-based. The advantages of using machine learning approach have been proven to achieve high accuracy in classification tasks and do not require complex linguistic resources [14].

Previous researches employed traditional word embedding method, namely Word2Vec, combined with CNN model to perform sentiment analysis in stock price prediction [12], [15]. Currently, there are various proposed modern LLMs (Large Language Model) that have been successfully applicable for NLP tasks. BERT (Bidirectional Encoder Representations from Transformers) is a kind of LLM which specifically designed for language understanding. BERT utilized Transformers architecture with bidirectional model. The bidirectional architecture enables BERT to understand the context of text in two directions (from left-to-right and right-to-left). Therefore, BERT is better than the language model with unidirectional architecture in understanding the context of the words. In addition, BERT using encoder part of Transformers which is better for natural language processing task that require deep understanding rather than generative tasks, such as sentiment analysis [16].

Therefore, this research proposed a stock price prediction model by utilizing sentiment analysis and historical stock price data. BERT is employed to perform sentiment analysis from tweets on X social media. Subsequently, the result of tweets sentiment analysis is integrated with historical stock price data to predict future stock price. The prediction of future actual stock price is performed by using BiLSTM. This research differs from the previous researches such as [17], [18] who utilized the sentiment analysis from social media post to predict the stock price movements rather than actual stock price data. It is expected that the use of sentiment analysis of public opinions on X as a feature is able to improve the accuracy of stock price predictions.

2. Method

This research performed stock price prediction using BiLSTM by utilizing sentiment analysis from X using BERT. Figure 1 explains the research stages for constructing and evaluating stock price prediction model. Sentiment analysis is performed on X post, known as tweets, to generate sentiment scores, which are then combined with historical stock price data. The combined data is used as input for the BiLSTM model to predict stock prices.

2.1. Data Collection from X

The data used to develop the sentiment analysis model was obtained from the X platform. The data collection process was carried out by utilizing Tweet-Harvest using the search keyword "IHSG". The research data covers the time range from September 22th, 2022 to May 11th, 2023. The total number of successfully retrieved data within that time range was 4,362 tweets. Some examples of the gathered tweets are displayed in Table 1.

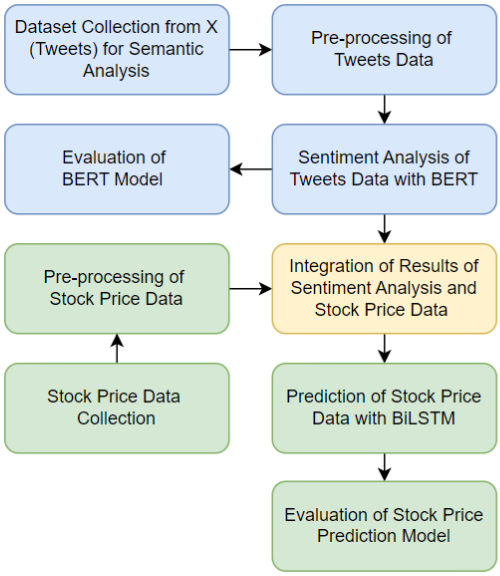

Fig. 1. Research Methodology.

Table 1. Example of Tweets on X related to IHSG stock price

Date	Tweets	Sentiment
2022-10-17	IHSG Senin Diprediksi Melemah Menanti Rilis Data Neraca Perdagangan RI	0
2022-12-10	IHSG Rabu masih lanjut melemah dipimpin saham sektor teknologi https://t.co/NgeUR9Kl4Z	0
2022-02-24	Jelang Akhir Pekan, IHSG Dibuka Menghijau https://t.co/vPicZorinI	1

The label of each tweet was obtained using two stages, namely manual labeling by researchers and validation of labeling results using VADER (Valence Aware Dictionary and Sentiment Reasoner). VADER produces a sentiment score by modifying the polarization value based on the emotional component in the text. Finally, each tweet was assigned with 0 for negative sentiment or 1 for positive sentiment. From a total of 4,362 tweets, there were 2,437 tweets with positive sentiment and 1,925 tweets with negative sentiment.

2.2. Preprocessing of Tweets

Prior to performing sentiment analysis from tweets data, text preprocessing was carried out to eliminate unnecessary information. First, emoticons in the text were removed using a special function that identifies and removes emoticons from the text. Next, URLs and usernames were removed using regular expressions.

The next step was the removal of punctuation marks and special characters that do not affect to the meaning of the text. This was done by using regular expressions to replace the punctuation marks and special characters with spaces. After that, repetitive characters in the text were removed, so that each character appears only twice. The next step was the removal of stop words, which are common words that do not have influences to the meaning of the text. Removing stop words allows the analysis to focus more on words that carry important meanings. Finally, the text was normalized to lowercase, so that all the posts have consistent lowercase format.

After data cleaning, the data preprocessing stage was performed as preparation for using the BERT model. One important step in preprocessing is text tokenization. In this stage, the text is divided into tokens that represent articulation units such as words or specific characters. Tokenization was done using a pre-initialized BERT tokenizer.

After the preprocessing stage of tweets from X was complete, the data was split into three parts, namely training data (60%), validation data (20%), and testing data (20%). Training and validation data is used to train and evaluate the BERT model.

2.3. Sentiment Analysis of Tweets

This research utilized BERT to perform sentiment analysis or classification from tweets on X. BERT is a kind of language representation that utilized Transformers architecture with bidirectional model. BERT was trained using 2,500 billion words from Wikipedia and 800 billion words from BooksCorpus [16]. This research used the model and bert-base-cased tokenizer from HuggingFace. This is the large model which has around 110 billion parameters. The text input in the BERT model is transformed into a form vector through three stages, namely token embeddings, segment embeddings, and position embeddings. The first stage will be converts text in a document into a dimensional vector representation. The second and third stages will produce one fixed token each for every existing token. In the second stage, the goal is to record information about which segment (sentence) each token belongs to. The third stage aims to provide an information model regarding the position of each word in a sentence [19].

2.4. Evaluation Metrics for Sentiment Analysis

Commonly used evaluation metrics in classification task was used to evaluate the performance of BERT model for sentiment classification in this research, namely accuracy, precision, recall, accuracy, and F1-score. The accuracy metric measures the proportion of correct classifications to the total classifications made. Precision is a metric that represents the proportion of correct positive classification predictions to the total positive classifications made. Recall is the proportion of correct positive classification predictions to the total actual positive classifications. F1-score is the average of recall and precision [20].

The output of the classification results by BERT can be categorized into four, True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN). TP and FP lead to positive labels, while TN and FN lead to negative labels. The prefix True means that the label prediction result is correct, while the prefix False means the classification result wrong. The value of TP, TN, FP, and FN is then used to calculate the accuracy, precision, recall, and F1-Score [20].

2.5. Stock Price Data Collection

Historical stock price data was obtained from the website id.investing.com. The IHSG stock price data for the period September 22th, 2022 to May 12th, 2023 was downloaded in Comma-Separated Values (CSV) format. The data includes information such as opening price, highest price, lowest price, closing price, volume, and adjusted closing price. An example of IHSG historical stock price data is displayed in Table 2.

Table 2. Example of historical IHSG stock price data.

Date	Open	Lowest	Highest	Close
2022-09-27	7127.50	7075,65	7133,41	7112,45
2022-09-28	7112.41	7073,47	7156,98	7077,03
2022-09-29	7077.11	7032,21	7135,50	7036,20

The selection of IHSG as the research object is due to its representation of overall stock market movements in the Indonesian stock exchange. Therefore, this research can provide relevant insights into general stock market trends. IHSG data is easily accessible and publicly available, facilitating the collection of historical data required for analysis and modeling.

2.6. Final Dataset

The final dataset in this research was a combination of sentiment score data and stock price data. The opening price, lowest price, highest price, and closing price of IHSG historical data from previous days were used as input features for prediction of closing stock price in a day after.

To integrate sentiment with the historical stock price data, each day's stock features were paired with a sentiment score with a one-day lag relative to the stock features. To handle multiple tweets within the same day, a daily averaged sentiment score was calculated, providing a single sentiment metric for each day. An example of the final data is displayed in Table 3. This data was processed and structured to fit as input for the BiLSTM model. The closing price was used as the target value for prediction.

Table 3. Example of final dataset.

Date	Open	Lowest	Highest	Close	Average Sentiment Score
27-09-2022	7.127,50	7.075,65	7.133,41	7.112,45	0.514
28-09-2022	7.112,41	7.073,47	7.156,98	7.077,20	0.000
29-09-2022	7.077,11	7.023,21	7.135,50	7.036,20	0.188

2.7. Preprocessing of Final Dataset

The data preprocessing stage was carried out by converting the values to dataset so that it is in the range 0 to 1. This is done by performing the min-max normalization procedure. Normalization is required to equalize the range of values for each attribute into the range 0-1.

2.8. Stock Price Prediction Using BiLSTM

This research employed BiLSTM to perform stock price prediction. The input for the BiLSTM model is combination of historical stock price data and the average of sentiment analysis score from previous day, reflecting a one-day lag between sentiment and stock data. BiLSTM is a variation of LSTM network which perform sequence processing in two directions, namely forward and backward direction [21].

LSTM is a special architecture of RNN which was created by Sepp Hochreiter and Jürgen Schmidhuber in 1997. LSTM trains the model using back propagation through time algorithm [22]. The LSTM cell, as shown in Figure 2, can remove or add information from the previous time-step in its memory (cell state) which is regulated via three special gates, namely an input gate, a forget gate, and an output gate [23].

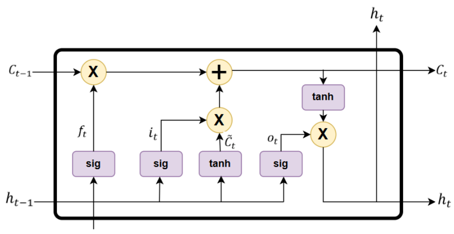

Fig. 2. LSTM cell [24]

The three gates (forget, input, output) and cell state has its own function to produce the last output for LSTM. Each gate and cell state has input weight (W), recurrent weight (U), and bias (b). The W_f, W_i, W_c, W_o are the input weight in the forget gate, input gate, cell state layer, and output gate, respectively. The U_f, U_i, U_c, U_o are the recurrent weight in the forget gate, input gate, cell state layer, and output gate, respectively. The b_f, b_i, b_c, b_o are the bias in the forget gate, input gate, cell state layer, and output gate, respectively.

The output from each gate is calculated based on the input in the current timestep x_t and the information from the previous hidden state h_{t-1}. Generally, LSTM involves three stages in calculating the value of the current hidden state h_t at timestep t. In the first stage, there is a forget gate in LSTM which has to decide which information to be forgotten or removed from the cell state. The forget gate uses sigmoid activation function so that the output from this gate ranges between [0,1] as shown by Equation (1) [24].

$$f_t = \text{sigmoid}(W_f.x_t + U_f.h_{t-1} + b_f) \quad (1)$$

The second stage is to decide which information that will be added to the cell state C_t. This calculation starts by calculating the output of input gate by using Equation (2). This input gate uses sigmoid activation function for deciding which value or state to be updated. Subsequently, the value of candidate cell state $\tilde{C}_t$ is calculated by using tanh activation function as shown by Equation (3). Therefore, the value of candidate cell state ranges between [-1,1]. Finally, the cell sate value is updated based on the values from the forget gate f_t, previous cell state C_{t-1}, input gate i_t, and candidate cell state $\tilde{C}_t$ as shown by Equation (4) [23], [24].

$$i_t = \text{sigmoid}(W_i.x_t + U_i.h_{t-1} + b_i) \quad (2)$$

$$\tilde{C}_t = tanh(W_c.x_t + U_c.h_{t-1} + b_c) \quad (3)$$

$$C_t = f_t.C_{t-1} + i_t.\tilde{C}_t \quad (4)$$

The third stage in LSTM cell is calculating the value of the current hidden state h_t which is the final output of LSTM cell. This is started by calculating the values of output gate by using sigmoid activation function as shown by Equation (5). The values of output gate decide which cell to be outputted. Then, the hidden state value is obtained from the pointwise multiplication between the output gate value o_t and the cell state value C_t that is activated by the tanh activation function asshown by Equation (6) [23], [24].

$$o_t = \text{sigmoid}(W_o.x_t + U_o.h_{t-1} + b_o) \quad (5)$$

$$h_t = o_t.\tanh(C_t) \quad (6)$$

Bidirectional LSTM (BiLSTM) is an optimized version of LSTM. As explained previously, the output of the LSTM in the current timestep will be calculated based on the output from the previous timestep. By using BiLSTM, the performance of the model increases due to the value of the current hidden state will be calculated based on the previous and subsequent information. BiLSTM network used two related LSTM layers, namely forward LSTM and backward LSTM as shown in Figure 3. The hidden state of forward LSTM layer is calculated based on the previous hidden state, while the hidden state of backward LSTM layer is calculated based on the subsequent hidden state [9], [21].

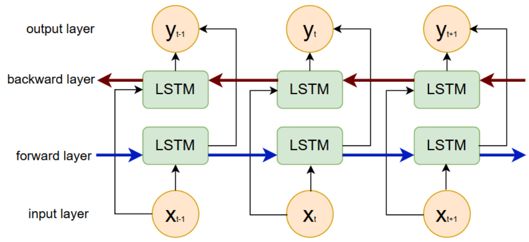

Fig. 3. BiLSTM network structure [9]

2.9. Evaluation of Stock Price Prediction Model

In this stage, the performance of the model in generating stock price predictions was evaluated. The model performance can be assessed by comparing the predicted prices with the actual prices. This research used MAPE (Mean Absolute Percentage Error) and RSME (Root Mean Squared Error) to calculate the error of the stock price prediction. The RMSE value is the result of the square root of the average of the squared differences actual value with predicted value as shown by Equation (7) where x_i is the actual value, $\hat{x}_i$ is the predicted value, and n is the number of evaluated data sample. The smaller the RMSE value means the value predictions are getting closer to the actual value. MAPE is the average of the absolute differences between actual values and predicted values as shown by Equation (8) [25].

$$RSME = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \hat{x}_i)^2}{n}} \quad (7)$$

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\frac{|(x_i - \hat{x}_i)^2|}{x_i} \quad (8)$$

3. Results and Discussion

3.1. Sentiment analysis of Tweets with BERT

This research tuned some hyperparameters of BERT model, namely learning rate and epoch, in order to obtain the best model for performing sentiment analysis of tweets from X. A proper learning rate value can solve local minimum problems. If the learning rate value is too high the model may skip the optimal solution, while if the learning is too small the model may be slower to reach the learning convergence. Apart from learning rate, the number of epochs or model training iterations also influence the performance of the model [26].

Table 4. The performance of BERT model using different learning rate and epoch.

Learning Rate	Epoch	Acc	F1-Score	Pre-cision	Recall
0.00001	1	0.647	0.410	0.823	0.273
0.00001	2	0.891	0.877	0.890	0.865
0.00001	3	0.922	0.916	0.884	0.952
0.00001	4	0.709	0.593	0.797	0.472
0.00001	5	0.551	0.513	0.495	0.315
0.00001	6	0.711	0.615	0.767	0.513
0.00001	7	0.715	0.627	0.760	0.533
0.00001	8	0.716	0.631	0.757	0.541
0.00001	9	0.712	0.631	0.744	0.548
0.00001	10	0.723	0.648	0.753	0.569
0.00001	11	0.724	0.652	0.751	0.577
0.00001	12	0.733	0.665	0.762	0.589
0.00001	13	0.729	0.663	0.749	0.594
0.00001	14	0.732	0.670	0.748	0.607
0.00001	15	0.727	0.665	0.742	0.602
0.001	1	0.771	0.734	0.697	0.775
0.001	2	0.551	0.591	0.575	0.511
0.001	3	0.810	0.790	0.781	0.801
0.001	4	0.809	0.775	0.821	0.735
0.001	5	0.817	0.787	0.822	0.755
0.001	6	0.824	0.799	0.816	0.783
0.001	7	0.822	0.793	0.832	0.758
0.001	8	0.824	0.796	0.827	0.768
0.001	9	0.828	0.801	0.836	0.768
0.001	10	0.832	0.806	0.834	0.781
0.001	11	0.830	0.807	0.826	0.788
0.001	12	0.828	0.804	0.825	0.783
0.001	13	0.842	0.827	0.811	0.844
0.001	14	0.841	0.816	0.847	0.788
0.001	15	0.841	0.816	0.848	0.786
0.1	1	0.551	0.051	0.497	0.442
0.1	2	0.449	0.620	0.449	1.000
0.1	3	0.551	0.523	0.425	0.514
0.1	4	0.507	0.560	0.468	0.699
0.1	5	0.565	0.059	0.522	0.306
0.1	6	0.551	0.047	0.497	0.425
0.1	7	0.701	0.569	0.808	0.439
0.1	8	0.691	0.574	0.752	0.464
0.1	9	0.582	0.145	0.886	0.080
0.1	10	0.691	0.716	0.609	0.867
0.1	11	0.708	0.597	0.784	0.482
0.1	12	0.580	0.128	0.931	0.069
0.1	13	0.565	0.059	0.611	0.031
0.1	14	0.706	0.651	0.696	0.612
0.1	15	0.639	0.454	0.708	0.334

Many studies have discussed the impact of adjustment hyperparameters to improve model quality, especially learning rate and number of epochs. This research tuned several learning rate values (0.1, 0.001, and 0.00001) and epoch (1, 2, 3, …, 15). The training process was conducted using batch size 8 and Adam optimizer.

Each hyperparameter combination (learning rate and epoch) was evaluated using the appropriate evaluation metrics, and the results are displayed in Table 4. The result of experiments shows that when using the highest learning rate of 0.1, the accuracy of model fluctuates among the epoch, as well as the precision, recall, and F1-score. However, the model reaches the highest accuracy of 0.708 when using learning rate of 0.1. When using learning rate of 0.001 the accuracy decreases at epoch 2, then increase gradually from epoch 3 and reach the highest accuracy of 0.842 in epoch 13. When using the lowest learning rate of 0.00001 the model increases drastically in epoch 2 and 3, the decrease gradually until epoch 5 and the increase again in the later epoch. The best accuracy of 0.922 was obtained when using the learning rate 0.00001 and epoch 3.

3.2. Stock Price Prediction Using BiLSTM

In order to observe the influence of sentiment analysis from tweets, this research performed the stock price prediction in two scenarios. The first scenario involved the result of tweet sentiment analysis as additional feature with the historical stock price data. The second scenario only performed prediction based on historical stock price data. Then, the result of the two scenarios can be compared to obtain the performance of sentiment analysis in predicting stock price data.

Scenario 1 – Stock Price Prediction with X Sentiment Analysis

After the best BERT model was obtained, the model was utilized to determine the sentiment of each collected tweet in previous day. The sentiment scores of the tweets were then averaged to create a single aggregated sentiment value, which was added as a feature for stock price prediction using the BiLSTM model. To perform stock price prediction using historical or time series data, the time step or window size must be properly determined. The window size represent how long (number of days) we involved the historical data in performing the prediction. If we use a window size of 3, it means that we use the three previous days to predict the stock price today. This research tuned several numbers of window size (3, 5, 8, 10, 15, 17, 20) to obtain the optimal model. Then, the model training was conducted using a BiLSTM layer with 50 units and batch size 32 in 150 epoch.

Table 5 shows the evaluation results using metrics RMSE and MAPE of various selected window sizes in scenario 1. Based on the results in Table 4, the lower error generally obtained when using the larger window size, such as 15, 17, and

19. However the error obtained were varied depend on the used window size. The lowest RSME and MAPE value were obtained when using window size 15.

Table 5. The performance of prediction model with different window size in scenario 1.

Window Size	RSME	MAPE
3	0.410	5.715
5	0.297	4.135
8	0.417	5.812
10	0.370	5.140
15	0.088	1.220
17	0.170	2.373
20	0.121	1.689

In addition, this research also performed experiment by stacking a BiLSTM layer in the previous model which used only a BiLSTM layer. Using more layers is possible to improve the capacity of model to learn more complex pattern. However, the risk of overfitting will be higher as the number of model parameter increase. Several number of nodes in the second BiLSTM layer was tuned to obtain the optimal model, such as 16, 32, 64, 128, and 256.

Table 6 shows the evaluation results using the window size 15 and different number of nodes in the second BiLSTM layer. Based on these results, it can be seen that the optimal number of nodes is 128 with the RMSE of 0.130 and MAPE of 1.806. By using the larger number of nodes, the error of model tends to decrease, except when using 256 nodes the RMSE and MAPE is higher than when using 128 nodes. This condition show that the performance of model increases when using larger number of nodes until certain number. However, when using too large number of nodes the error may increase as the model will have more network parameters which is possible to increase the risk overfitting.

Table 6. The performance of prediction model with different number of nodes in the second BiLSTM in scenario 1.

Number of Nodes	RSME	MAPE
16	0.150	2.090
32	0.146	2.036
64	0.137	1.912
128	0.130	1.806
256	0.135	1.880

The RMSE and MAPE of best model when using two BiLSTM layer are higher compared to the RMSE and MAPE of model when using only a BiLSTM layer. In this case using a BiLSTM layer is sufficient to learn the pattern of training data, while adding more layer may increase the complexity of model.

The combination of sentiment scores and closing stock prices reveals a correlation between the sentiment in social media and stock price movements, as shown in Table 7 For instance, a decrease in sentiment from 0.620 to 0.404 on March 6th to March 7th corresponds with a fall in the stock price from 6807.00 to 6766, while an increase in the sentiment score to 0.500 in the next day corresponds with a rise in the stock price to 6776.37. This demonstrates the potential influence of public sentiment on stock prices, which can be further analyzed to determine predictive patterns.

Table 7. Combination of sentiment score and closing stock price.

Date	Sentiment Score	Closing Price
2023-03-01	0.649	6844.93
2023-03-02	0.653	6857.41
...	...	...
2023-03-06	0.620	6807.00
2023-03-07	0.404	6766.75
2023-03-08	0.500	6776.37
2023-03-09	0.690	6799.80
...	...	...

Scenario 2 – Stock Price Prediction without Sentiment Analysis

The second scenario performed prediction based on only historical stock price data. The configuration of BiLSTM network for training was same as used in the scenario 1. The first experiment was performed to search for the best window size to predict the stock price. Table 8 shows the evaluation results using metrics RMSE and MAPE of various selected window_sizes in scenario 2. As the result in Table 6, the RMSE and MAPE of the model are getting lower when using the higher window size untill window size of 15, but if the window size continues to be increased, the RMSE and MAPE of the model are getting higher. The best window size obtained in this scenario is same as the best window size in the scenario 1.

Table 8. The performance of prediction model with different window size in scenario 2.

Window Size	RSME	MAPE
3	0.423	5.895
5	0.373	5.190
8	0.361	5.027
10	0.488	6.799
15	0.153	2.135
17	0.235	3.267
20	0.196	2.725

Subsequently, this scenario also performed experiment by stacking a BiLSTM layer in the previous model as performed in the scenario 1. Table 9 shows the evaluation results using the window size 15 and different number of nodes in the second BiLSTM layer. Based on this results, it can be seen that the optimal number of nodes is 32 with the RMSE of 0.130 and MAPE of 1.815. The RMSE and MAPE of the model when using 32 nodes are lower than when using 16 nodes.

However, if the number of nodes continues to be increased, the RMSE and MAPE of the model are getting higher. This condition implay that in this case using too large number of nodes is not suitable since may increase the number of network parameters and also increase the risk of overfitting.

Table 9. The performance of prediction model with different number of nodes in the second BiLSTM in scenario 2.

Number of Nodes	RSME	MAPE
16	0.156	2.176
32	0.130	1.815
64	0.133	1.853
128	0.134	1.867
256	0.138	1.922

The RMSE and MAPE of the best model when using two BiLSMT layer are lower compared to the RMSE and MAPE of the best model when using only a BiLSTM layer. In this case, adding the second layer can improve the model's ability to learn data patterns better.

Comparison of Scenario 1 and Scenario 2

The comparison of the best model in scenario 1 (based on sentiment analysis and historical stock price data) and scenario 2 (based on historical stock price data) can be seen in Table 10. The results of experiment show that the RMSE and MAPE of the best model in scenario 1 that utilized tweets sentiment analysis is decreased by 0.042 and 0.595, resepectively compared to the best prediction model without sentiment analysis.

Table 10. The performance of prediction model with different number of nodes in the second BiLSTM in scenario 1.

Scenario	Window size	2nd BiLSTM layer	RSME	MAPE
1	15	No	0.088	1.220
2	15	Yes (32 nodes)	0.130	1.815

4. Conclusion

The proposed study presents a framework for stock price prediction that incorporates historical stock data and sentiment analysis related to IHSG stock price from X. The sentiment analysis is performed using BERT model to calculate sentiment polarity indicating the positivity or negativity of the tweets related to IHSG topic from X social media. The study examines the effect of utilizing sentiment analysis from X on stock price predictions over different time step or window size, including 5 days, 10 days, and 15 days. The performance of model was then evaluated in term of RMSE and MAPE. The results suggest that using sentiment analysis from X with BERT in conjunction with BiLSTM succeeded to improve the performance of the stock price prediction model. The RMSE and MAPE of the proposed model that utilized Tweets sentiment analysis is decreased by 0.042 and 0.595, resepectively, compared to the RMSE and MAPE of the prediction model without sentiment analysis.
These results demonstrate a significant correlation between stock price and tweets from X. Future studies can explore more specific feature, including social media posts, political news, and geopolitical events, along with financial news, which is possible to enhance the performance of stock price prediction model.

Acknowledgement

This work was supported by the Faculty of Science and Mathematics, Universitas Diponegoro under the scheme of Student Research in 2023.

References

[1] Otoritas Jasa Keuangan, *Buku Seri Literasi Keuangan Tingkat Perguruan Tinggi - Pasar Modal*. Jakarta: Otoritas Jasa Keuangan, 2016.

[2] D. S. Sampurna, “Analisis Pengaruh Faktor-Faktor Ekonomi Makro Terhadap IHSG Di Bursa Efek Indonesia (BEI),” *Jurnal STEI Ekonomi*, vol. 25, no. 1, pp. 54–73, 2016.

[3] A. Ashta and H. Herrmann, “Artificial intelligence and fintech: An overview of opportunities and risks for banking, investments, and microfinance,” *Strategic Change*, vol. 30, no. 3, pp. 211–222, May 2021, doi: 10.1002/jsc.2404.

[4] B. Gülmez, “Stock price prediction with optimized deep LSTM network with artificial rabbits optimization algorithm,” *Expert Syst Appl*, vol. 227, no. 120346, Oct. 2023, doi: 10.1016/j.eswa.2023.120346.

[5] M. Tan, C. dos Santos, B. Xiang, and B. Zhou, “LSTM-based Deep Learning Models for Non-factoid Answer Selection,” in *ICLR*, Nov. 2016. [Online]. Available: http://arxiv.org/abs/1511.04108

[6] A. C. Ozdemir, K. Buluş, and K. Zor, “Medium- to long-term nickel price forecasting using LSTM and GRU networks,” *Resources Policy*, vol. 78, no. 102906, Sep. 2022, doi: 10.1016/j.resourpol.2022.102906.

[7] D. Yao and K. Yan, “Time series forecasting of stock market indices based on DLWR-LSTM model,” *Financ Res Lett*, vol. 68, no. 105821, Oct. 2024, doi: 10.1016/j.frl.2024.105821.

[8] A. Q. Md *et al.*, “Novel optimization approach for stock price forecasting using multi-layered sequential LSTM,” *Appl Soft Comput*, vol. 134, no. 109830, Feb. 2023, doi: 10.1016/j.asoc.2022.109830.

[9] A. Kulshrestha, V. Krishnaswamy, and M. Sharma, “Bayesian BILSTM approach for tourism demand forecasting,” *Ann Tour Res*, vol. 83, no. 102925, Jul. 2020, doi: 10.1016/j.annals.2020.102925.

[10] B. Bhattacharjee, R. Kumar, and A. Senthilkumar, “Unidirectional and bidirectional LSTM models for edge weight predictions in dynamic cross-market equity networks,” *International Review of Financial Analysis*, vol. 84, no. 102384, Nov. 2022, doi: 10.1016/j.irfa.2022.102384.

[11] M. P. Cristescu, D. A. Mara, R. A. Nerișanu, L. C. Culda, and I. Maniu, “Analyzing the Impact of Financial News Sentiments on Stock Prices—A Wavelet Correlation,”

Mathematics, vol. 11, no. 23, Dec. 2023, doi: 10.3390/math11234830.

[12] N. Jing, Z. Wu, and H. Wang, "A hybrid model integrating deep learning with investor sentiment analysis for stock price prediction," *Expert Syst Appl*, vol. 178, no. 115019, Sep. 2021, doi: 10.1016/j.eswa.2021.115019.

[13] Q. Liu, W. S. Lee, M. Huang, and Q. Wu, "Synergy between stock prices and investor sentiment in social media," *Borsa Istanbul Review*, vol. 23, no. 1, pp. 76–92, Jan. 2023, doi: 10.1016/j.bir.2022.09.006.

[14] M. D. Devika, C. Sunitha, and A. Ganesh, "Sentiment Analysis: A Comparative Study on Different Approaches," *Procedia Comput Sci*, vol. 87, pp. 44–49, 2016, doi: 10.1016/j.procs.2016.05.124.

[15] Q. Liu, W. S. Lee, M. Huang, and Q. Wu, "Synergy between stock prices and investor sentiment in social media," *Borsa Istanbul Review*, vol. 23, no. 1, pp. 76–92, Jan. 2023, doi: 10.1016/j.bir.2022.09.006.

[16] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in *Proceedings of the 2019 Conference of the North*, Stroudsburg, PA, USA: Association for Computational Linguistics, 2019, pp. 4171–4186. doi: 10.18653/v1/N19-1423.

[17] A. Derakhshan and H. Beigy, "Sentiment analysis on stock social media for stock price movement prediction," *Eng Appl Artif Intell*, vol. 85, pp. 569–578, Oct. 2019, doi: 10.1016/j.engappai.2019.07.002.

[18] J. X. Liu, J. S. Leu, and S. Holst, "Stock price movement prediction based on Stocktwits investor sentiment using FinBERT and ensemble SVM," *PeerJ Comput Sci*, vol. 9, 2023, doi: 10.7717/peerj-cs.1403.

[19] S. Khatoon *et al.*, "Development of Social Media Analytics System for Emergency Event Detection and Crisis Management," *Computers, Materials & Continua*, vol. 68, no. 3, pp. 3079–3100, 2021, doi: 10.32604/cmc.2021.017371.

[20] J. Han and M. Kamber, *Data Mining: Concepts and Techniques Second Edition*. San Farnsisco: Elsevier Inc., 2006.

[21] N. Zrira, A. Kamal-Idrissi, R. Farssi, and H. A. Khan, "Time series prediction of sea surface temperature based on BiLSTM model with attention mechanism," *J Sea Res*, vol. 198, Apr. 2024, doi: 10.1016/j.seares.2024.102472.

[22] A. Mittal, V. Dhiman, A. Singh, and C. Prakash, "Short-Term Bitcoin Price Fluctuation Prediction Using Social Media and Web Search Data," *2019 12th International Conference on Contemporary Computing, IC3 2019*, pp. 1–6, 2019, doi: 10.1109/IC3.2019.8844899.

[23] F. Liu and C. Liang, "Short-term power load forecasting based on AC-BiLSTM model," *Energy Reports*, vol. 11, pp. 1570–1579, Jun. 2024, doi: 10.1016/j.egyr.2024.01.026.

[24] Y. Wu, M. Yuan, S. Dong, L. Lin, and Y. Liu, "Remaining useful life estimation of engineered systems using vanilla LSTM neural networks," *Neurocomputing*, vol. 275, pp. 167–179, Jan. 2018, doi: 10.1016/j.neucom.2017.05.063.

[25] A. V Tatachar, "Comparative Assessment of Regression Models Based On Model Evaluation Metrics," *International Research Journal of Engineering and Technology*, vol. 8, no. 9, 2021, [Online]. Available: www.irjet.net

[26] J. Krohn, G. Beyleveld, and A. Bassens, *Deep Learning Illustrated: A Visual, Interactive Guide to Artificial Intelligence (Addison-Wesley Data & Analytics Series) 1st Edition*, 1st ed. Boston: Pearson Addison-Wesley, 2020.